

Introductie tot Data Science & Machine learning

Leer data beter te analyseren en leg een basis voor betrouwbare voorspellingen

U bent bekend met het werken met data. Nu wordt van u verwacht om inzichten te halen uit grote hoeveelheden data. Om betrouwbare analyses te maken heeft u de juiste statistische technieken en inzichten nodig. Met de opleiding Introductie tot Data Science & Machine learning leert u zelf aan de hand van R op praktische wijze hoe u in de verzamelde gegevens datapatronen kunt ontdekken en voorspellingen kunt doen. Als u niet zelf aan de slag gaat met data en toch een goed begrip wil hebben van de methoden, dan is deze opleiding ook een goede keuze. U bent met de opgedane kennis een goede gesprekspartner voor data-analisten en management.

Organisaties verzamelen steeds meer data. Vanuit de business is er behoefte om op gefundeerde wijze gebruik te maken van deze gegevens. Veel informatie is echter ongestructureerd terwijl de vragen uit de organisatie vaak slecht gearticuleerd zijn: wat wil men weten? De opleiding Introductie tot data science en machine learning leert u welke methoden u kunt gebruiken voor verschillende onderzoeksvragen. Daarnaast leert u de kwaliteit en betekenis van de resultaten te beoordelen. Aan de hand van R gaat u zelf aan de slag met het uitvoeren van analyses. U leert op praktische wijze zelf met verschillende data analyse methoden te werken en deze toe te passen.

Verschillende methoden

De opleiding biedt een gedegen inleiding op een aantal breed inzetbare data-analyse methoden:

- Clusteranalyse
- Dimensie reductie en data visualisatie
- Beslissingsbomen en ensembles van beslissingsbomen
- Support vector machines (SVM)

De theorie achter deze methoden komt uitgebreid aan bod.

Wanneer

Woensdag 4, 11 december 2019 en 15, 22 januari 2020.

Prijs

De investering bedraagt: € 2725 (vrijgesteld van BTW) Inbegrepen zijn dagarrangement en cursusmateriaal

Waar

Novotel Brainpark Rotterdam

Direct inschrijven

www.erasmusacademie.nl/introductiedatascience

Aantal deelnemers

Het maximum aantal deelnemers bedraagt 20

Certificaat

Op de slotdag van de opleiding ontvangt u als bewijs van deelname een certificaat.

Meer informatie

Neem contact op met Miranda Smit.
Adviseur opleidingen T: 010-4088687,
E: smit@erasmusacademie.nl

Zelf aan de slag in R

De opleiding besteedt veel aandacht aan het zelf kunnen werken met deze methoden. Aan de hand van datasets oefent u met de verschillende statistische analyses in het programma R. U ervaart wat de gebruiksmogelijkheden zijn in de praktijk.

Toepassingsgericht

Naast de theorie en het uitvoeren van de analyses is er ruime aandacht voor de toepassing. U kunt daarbij denken aan vragen als: wat doet de methode (functie), wanneer is deze nuttig (doel), welke vragen kan ik ermee beantwoorden (resultaat) en waar moet ik op letten bij het interpreteren van de resultaten (risico's)?

Resultaten

- U maakt kennis met verschillende data-analyse methoden
- U kunt vaststellen welke methode(n) geschikt zijn om uw onderzoeksvraag te beantwoorden;
- U leert zelf verschillende soorten analyses in R uit te voeren;
- U kunt de kwaliteit en betekenis van de resultaten beoordelen;
- U bent door beter begrip van termen en methoden een goede gesprekspartner voor management en data team.

Programma

Dag 1: woensdag 4 december 2019

U leert op een gestructureerde manier kijken naar big data en het formuleren van relevante probleemstellingen. U ontvangt uitleg over het werken met R. U ontdekt wat dimensiereductie en data-visualisatie is en hoe u dit kunt toepassen.

Dag 2: woensdag 11 december 2019

U maakt kennis met cluster analyse: wat doet het, welke keuzes zijn er nodig (en hoe komen deze keuzes tot stand). U ontdekt hoe en wanneer u clusteranalyse kunt uitvoeren en u leert resultaten te interpreteren.

Dag 3: woensdag 15 januari 2020

Op dag 3 worden beslissingsbomen behandeld. U ontdekt de sterktes en zwaktes van beslissingsbomen. U ziet in wat wel en niet vergeleken kan worden in termen van validatie en voorspelmethoden en begrijpt welke kwaliteitsmaten niet en, vooral, wel zinvol zijn.

Dag 4: woensdag 22 januari 2020

Op deze dag komen twee geavanceerde machine learning methoden aan bod die bekend staan als zogenaamde "black-box" methoden: Random Forests en Support Vector Machines (SVM). U leert wat de methodes doen, waarom/wanneer ze goed werken, hoe u de methoden zelf optimaal kunt trainen en hoe u de resultaten kunt duiden.

Vorbereiding

U wordt geacht in het bezit te zijn van een laptop waarop het freeware programma R is gedownload en waarmee u de statistische berekeningen kunt uitvoeren. Het freeware programma R is beschikbaar voor Windows, Linux en Mac en via www.r-project.org gratis te downloaden. Voorkennis van R is niet noodzakelijk.

Docent

Michel van de Velden

Michel van de Velden is universitair hoofddocent (UHD) Statistiek bij het Econometrisch Instituut van de Erasmus School of Economics. Michel van de Velden promoveerde in 2000 aan de Universiteit van Amsterdam. Na zijn promotie verwierf hij een Grotius Grant en een Marie Curie Fellowship waarmee hij onderzoek deed bij de rijksuniversiteit Groningen en de Universitat Pompeu Fabra te Barcelona.

Voor wie?

U bent een beginnende data analist of data manager. Uw organisatie wil stappen zetten op het gebied van Big data, statistische analyses en voorspellen. U wilt zelf bewerkingen kunnen uitvoeren op basis van grote hoeveelheden data, dan wel uw team inhoudelijk kunnen aansturen. U bent praktisch ingesteld, resultaatgericht en wil de uitkomsten van uw analyses graag nuttig gebruiken voor de business. Uw basiskennis van statistiek is goed en u vraagt zich af hoe u data moet ordenen, welke methoden u het beste kunt toepassen en hoe u komt tot betrouwbare voorspellingen.

Liggen uw leerdoelen meer op het gebied van voorspellen met data in plaats van op het analyseren dan is de opleiding Data Science en Predictive Machine learning voor u een goede keuze.

Wij zien uw inschrijving graag tegemoet:

www.erasmusacademie.nl/introductiedatascience

"Erg leerzame training waarbij ogenschijnlijk lastige materie in duidelijk stappen wordt uitgelegd. Fijne docent."

"Ik vond het heel knap hoe het materiaal en de docent ervoor zorgden dat de deelnemers het goed konden volgen en er van hebben geleerd, ondanks de vele verschillende achtergronden van de deelnemers in werkterrein, leeftijd en ervaring met statistiek en R."